

## STIC Biotechnology Systems Branch

### RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/797,553B  
Source: 1fw0  
Date Processed by STIC: 3/10/05

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.2.2 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):  
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/24/05

## Raw Sequence Listing Error Summary

### ERROR DETECTED

### SUGGESTED CORRECTION

SERIAL NUMBER: 10/797,553B

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☐ Wrapped Nucleics  
Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 ☒ Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 ☒ Misaligned Amino  
Numbering The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 ☐ Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 ☒ Variable Length Sequence(s) 56 contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 ☐ PatentIn' 2.0  
"bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) \_\_\_\_\_. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 ☐ Skipped Sequences  
(OLD RULES) Sequence(s) \_\_\_\_ missing. If intentional, please insert the following lines for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
This sequence is intentionally skipped  
  
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 ☐ Skipped Sequences  
(NEW RULES) Sequence(s) \_\_\_\_ missing. If intentional, please insert the following lines for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 9 ☐ Use of n's or Xaa's  
(NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ☐ Invalid <213>  
Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ☐ Use of <220> Sequence(s) \_\_\_\_ missing the <220> "Feature" and associated numeric identifiers and responses.  
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 ☐ PatentIn 2.0  
"bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 ☐ Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:09

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

3 <110> APPLICANT: Moyle, William R.  
 4 Xing, Yongna  
 6 <120> TITLE OF INVENTION: Protein Knobs  
 8 <130> FILE REFERENCE: 268/279-RWJ-01-40  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/797,553B  
 C--> 11 <141> CURRENT FILING DATE: 2004-03-10  
 E--> 13 <160> NUMBER OF SEQ ID NOS: 5666 (p.9)  
 15 <170> SOFTWARE: PatentIn version 3.1

*submitted file was  
 not in ASCII text  
 format. Do NOT submit  
 a PatentIn ".prj" file;  
 it is a  
 Does Not Comply work-in-  
 corrected Diskette program  
 Needer file.  
 see pp 1-15*

*submit a  
 generated  
 sequence  
 listing.*

## ERRORED SEQUENCES

755 <210> SEQ ID NO: 24  
 756 <211> LENGTH: 92  
 757 <212> TYPE: PR1  
 758 <213> ORGANISM: Artificial Sequence  
 760 <220> FEATURE:  
 761 <223> OTHER INFORMATION: hCG alpha-subunit with Cys substituted for Lys51  
 763 <400> SEQUENCE: 24  
 765 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
 766 1 5 10 15  
 769 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
 770 20 25 30  
 773 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
 774 35 40 45  
 777 Val Cys Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
 778 50 55 60  
 781 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
 782 65 70 75 80  
 785

*what is  
 this?*

Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu  
 E--> 786 1 5 10 15  
 789 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
 E--> 790 20 25 30  
 793 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
 E--> 794 35 40 45  
 797 Val Gln Cys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
 E--> 798 50 55 60  
 801 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
 E--> 802 65 70 75 80  
 805 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
 E--> 806 85 90  
 1183 <210> SEQ ID NO: 36  
 1184 <211> LENGTH: 145

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:09

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

```

1185 <212> TYPE: PRT
1186 <213> ORGANISM: Homo sapiens
1188 <400> SEQUENCE: 36
1190 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
1191 1 5 10 15
1194 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
1195 20 25 30
1198 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
1199 35 40 45
1202 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
1203 50 55 60
1206 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Asn Pro Val Val
1207 65 70 75 80
1210 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser
1211 85 90 95
1214 Thr Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
1215 100 105 110
1218 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
1219 115 120 125
1222 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln
E--> 1223 130 135 140
1226 <210> SEQ ID NO: 37
1227 <211> LENGTH: 145
1228 <212> TYPE: PRT
1229 <213> ORGANISM: Artificial Sequence
1231 <220> FEATURE:
1232 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Ser138
1234 <400> SEQUENCE: 37
1236 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
1237 1 5 10 15
1240 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
1241 20 25 30
1244 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
1245 35 40 45
1248 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
1249 50 55 60
1252 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Asn Pro Val Val
1253 65 70 75 80
1256 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser
1257 85 90 95
1260 Thr Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
1261 100 105 110
1264 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
1265 115 120 125
1268 Pro Ser Pro Ser Arg Leu Pro Gly Pro Cys Asp Thr Pro Ile Leu Pro Gln
E--> 1269 130 135 140
1272 <210> SEQ ID NO: 38
1273 <211> LENGTH: 145
1274 <212> TYPE: PRT

```

insert a hard  
return. A maximum  
of 16 amino acids  
per line.

145 ← insert

insert hard  
return - per Sequen  
Rules a maximum of  
16 amino acids per  
line  
145 ← insert number line

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:09

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

1275 &lt;213&gt; ORGANISM: Artificial Sequence

1277 &lt;220&gt; FEATURE:

1278 <223> OTHER INFORMATION: hCG beta-subunit residues 101-114 were replaced with their  
hFSH b

1279 eta-subunit counterparts, namely hFSH beta-subunit residues 95-10

1280 8

1282 &lt;400&gt; SEQUENCE: 38

1284 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu

1285 1 5 10 15

1288 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr

1289 20 25 30

1292 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val

1293 35 40 45

1296 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe

1297 50 55 60

1300 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val

1301 65 70 75 80

1304 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser

1305 85 90 95

1308 Thr Thr Asp Cys Thr Val Arg Gly Leu Gly Pro Ser Tyr Cys Ser Phe

1309 100 105 110

1312 Gly Glu Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu

1313 115 120 125

1316 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln  
E--> 1317 130 135 140 145

1320 &lt;210&gt; SEQ ID NO: 39

1321 &lt;211&gt; LENGTH: 145

1322 &lt;212&gt; TYPE: PRT

1323 &lt;213&gt; ORGANISM: Artificial Sequence

1325 &lt;220&gt; FEATURE:

1326 <223> OTHER INFORMATION: hCG beta-subunit residues 101-114 were replaced with their  
hFSH b

1327 eta-subunit counterparts, namely hFSH beta-subunit residues 95-10

1328 8, and Serine38 in the beta-subunit carboxyterminus of this

1329 analog was replaced with Cys

1331 &lt;400&gt; SEQUENCE: 39

1333 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu

1334 1 5 10 15

1337 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr

1338 20 25 30

1341 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val

1342 35 40 45

1345 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe

1346 50 55 60

1349 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val

1350 65 70 75 80

1353 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser

1354 85 90 95

1357 Thr Thr Asp Cys Thr Val Arg Gly Leu Gly Pro Ser Tyr Cys Ser Phe

1358 100 105 110

1361 Gly Glu Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:09

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

```

1362          115          120          125
1365 Pro Ser Pro Ser Arg Leu Pro Gly Pro Cys Asp Thr Pro Ile Leu Pro Gln
E--> 1366      130          135          140      145
1729 <210> SEQ ID NO: 45
1730 <211> LENGTH: 125
1731 <212> TYPE: PRT
1732 <213> ORGANISM: Artificial Sequence
1734 <220> FEATURE:
1735 <223> OTHER INFORMATION: hCgbeta,delta116-135,S138C
1737 <400> SEQUENCE: 45
1739 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
1740 1          5          10          15
1743 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
1744          20          25          30
1747 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
1748          35          40          45
1751 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
1752          50          55          60
1755 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Asn Pro Val Val
1756 65          70          75          80
1759 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser
1760          85          90          95
1763 Thr Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
1764          100         105         110
1767 Pro Arg Phe Gly Pro Cys Asp Thr Pro Ile Leu Pro Gln
E--> 1768      115          120      125 ←
1863 <210> SEQ ID NO: 48
1864 <211> LENGTH: 140
1865 <212> TYPE: PRT
1866 <213> ORGANISM: Artificial Sequence
1868 <220> FEATURE:
1869 <223> OTHER INFORMATION: hCgbeta,delta131-135,S138C
1871 <400> SEQUENCE: 48
1873 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
1874 1          5          10          15
1877 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
1878          20          25          30
1881 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
1882          35          40          45
1885 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
1886          50          55          60
1889 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Asn Pro Val Val
1890 65          70          75          80
1893 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser
1894          85          90          95
1897 Thr Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
1898          100         105         110
1901 Pro Arg Phe Gln Asp Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
1902          115          120          125

```

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:10

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

1905 Pro Ser Gly Pro Cys Asp Thr Pro Ile Leu Pro Gln  
E--> 1906 130 135 1406  
2143 <210> SEQ ID NO: 56  
2144 <211> LENGTH: 10  
2145 <212> TYPE: PRT  
2146 <213> ORGANISM: Artificial Sequence *this needs explanation in 2207-2237 section*  
2148 <220> FEATURE:  
2149 <223> OTHER INFORMATION: Xl-Asp-Asp-Asp-Asp-Lys-Ser-Ym-Cys-Zn, where X, Y, and Z refer to  
2150 any tail portion amino acids and l, m, and n refer to the lengths *see item 5 on Error summary sheet*  
2151 of the tail portion amino acids  
2153 <220> FEATURE:  
2154 <221> NAME/KEY: MISC\_FEATURE  
2155 <223> OTHER INFORMATION: Xaa refers to any tail portion amino acids and n refers to the  
2156 lengths of the tail portion amino acids *n's are not permitted in the sequence*  
2160 <400> SEQUENCE: 56  
E--> 2162 Xaa n Asp Asp Asp Asp Lys Ser Xaa n Cys Xaa n *misaligned numbers, see item 3 on Error summary sheet*  
E--> 2163 1 5 10 10  
2167 <210> SEQ ID NO: 57  
2168 <211> LENGTH: 92 *Artificial*  
2169 <212> TYPE: PRT  
C--> 2170 <213> ORGANISM: Artificial Sequence  
2172 <220> FEATURE:  
2173 <223> OTHER INFORMATION: An hCG truncated (-subunit analog fused to the hCG alpha-carboxyterminus  
2175 <400> SEQUENCE: 57  
2177 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
2178 1 5 10 15  
2180 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
2181 20 25 30  
2183 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
2184 35 40 45  
2186 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
2187 50 55 60  
2189 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
2190 65 70 75 80  
2192 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser Asp Asp Pro Arg  
E--> 2193 85 90 95 90 95 *misaligned numbers*  
2195 Phe Gly Pro Cys Asp Thr Pro Ile Leu Pro Gln  
E--> 2196 100 105 100 105  
2198 <210> SEQ ID NO: 58  
2199 <211> LENGTH: 145  
2200 <212> TYPE: PRT  
2201 <213> ORGANISM: Artificial Sequence  
2203 <220> FEATURE:  
2204 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Arg94  
2206 <400> SEQUENCE: 58  
2208 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu  
2209 1 5 10 15  
2212 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr  
2213 20 25 30

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:10

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

```

2216 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
2217      35      40      45
2220 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
2221      50      55      60
2224 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val
2225      65      70      75      80
2228 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Cys Arg Ser
2229      85      90      95
2232 Thr Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
2233      100     105     110
2236 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
2237      115     120     125
2240 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln
E--> 2241      130     135     140
2244 <210> SEQ ID NO: 59
2245 <211> LENGTH: 145
2246 <212> TYPE: PRT
2247 <213> ORGANISM: Artificial Sequence
2249 <220> FEATURE:
2250 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Arg95
2252 <400> SEQUENCE: 59
2254 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
2255      1      5      10      15
2258 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
2259      20      25      30
2262 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
2263      35      40      45
2266 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
2267      50      55      60
2270 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val
2271      65      70      75      80
2274 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Cys Ser
2275      85      90      95
2278 Thr Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
2279      100     105     110
2282 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
2283      115     120     125
2286 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln
E--> 2287      130     135     140
2290 <210> SEQ ID NO: 60
2291 <211> LENGTH: 145
2292 <212> TYPE: PRT
2293 <213> ORGANISM: Artificial Sequence
2295 <220> FEATURE:
2296 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Ser96
2298 <400> SEQUENCE: 60
2300 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
2301      1      5      10      15
2304 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr

```

*hard return*

*hard return*

*P. 7*



## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:10

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

```

2305          20          25          30
2308 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
2309          35          40          45
2312 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
2313          50          55          60
2316 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val
2317 65          70          75          80
2320 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Cys
2321          85          90          95
2324 Thr Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
2325          100         105         110
2328 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
2329          115         120         125
2332 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln
E--> 2333          130         135         140
2335 <210> SEQ ID NO: 61
2336 <211> LENGTH: 145
2337 <212> TYPE: PRT
2338 <213> ORGANISM: Artificial Sequence
2340 <220> FEATURE:
2341 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Thr97
2343 <400> SEQUENCE: 61
2345 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
2346 1          5          10          15
2349 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
2350          20          25          30
2353 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
2354          35          40          45
2357 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
2358          50          55          60
2361 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val
2362 65          70          75          80
2365 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser
2366          85          90          95
2369 Cys Thr Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
2370          100         105         110
2373 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
2374          115         120         125
2377 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln
E--> 2378          130         135         140
2380 <210> SEQ ID NO: 62
2381 <211> LENGTH: 145
2382 <212> TYPE: PRT
2383 <213> ORGANISM: Artificial Sequence
2385 <220> FEATURE:
2386 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Thr98
2388 <400> SEQUENCE: 62
2390 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
2391 1          5          10          15

```

hard return

hard return

145 ← insert

P.8

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:10

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

```

2394 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
2395                20                25                30
2398 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
2399                35                40                45
2402 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
2403                50                55                60
2406 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val
2407 65                70                75                80
2410 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser
2411                85                90                95
2414 Thr Cys Asp Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
2415                100               105               110
2418 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
2419                115               120               125
2422 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln
E--> 2423                130               135               140
2425 <210> SEQ ID NO: 63
2426 <211> LENGTH: 145
2427 <212> TYPE: PRT
2428 <213> ORGANISM: Artificial Sequence
2430 <220> FEATURE:
2431 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Asp99
2433 <400> SEQUENCE: 63
2435 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu
2436 1                5                10                15
2439 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr
2440                20                25                30
2443 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val
2444                35                40                45
2447 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe
2448                50                55                60
2451 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val
2452 65                70                75                80
2455 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Ser
2456                85                90                95
2459 Thr Thr Cys Cys Gly Gly Pro Lys Asp His Pro Leu Thr Cys Asp Asp
2460                100               105               110
2463 Pro Arg Phe Gln Asp Ser Ser Ser Ser Lys Ala Pro Pro Pro Ser Leu
2464                115               120               125
2467 Pro Ser Pro Ser Arg Leu Pro Gly Pro Ser Asp Thr Pro Ile Leu Pro Gln
E--> 2468                130               135               140
2470 <210> SEQ ID NO: 64
2471 <211> LENGTH: 95
2472 <212> TYPE: PRT
C--> 2473 <213> ORGANISM: Artificial Sequence
2475 <220> FEATURE:
2476 <223> OTHER INFORMATION: An hCG alpha-subunit analog with Gly-Gly-Cys at its
carboxyterminus
2478 <400> SEQUENCE: 64
2480 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro

```

hard return

hard return

space

P.9

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/797,553B

DATE: 03/10/2005  
TIME: 09:59:10

Input Set : A:\Sequen-3.prj  
Output Set: N:\CRF4\03102005\J797553B.raw

2481 1 5 10 15  
2483 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
2484 20 25 30  
2486 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
2487 35 40 45  
2489 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
2490 50 55 60  
2492 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
2493 65 70 75 80  
2495 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser Gly Gly Cys  
E--> 2496 ~~86 90~~ 85 95 90 95  
2499 <210> SEQ ID NO: 65  
2500 <211> LENGTH: 92  
2501 <212> TYPE: PRT  
C--> 2502 <213> ORGANISM: Artificial Sequence  
2504 <220> FEATURE:  
2505 <223> OTHER INFORMATION: An hCG alpha-subunit analog with Asp in place of Asn52 and  
Cys in place of Ser92  
2507 <400> SEQUENCE: 65  
2509 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
2510 1 5 10 15  
2512 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
2513 20 25 30  
2515 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
2516 35 40 45  
2518 Val Gln Lys Asp Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
2519 50 55 60  
2521 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
2522 65 70 75 80  
2524 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys (Ser) fix nos.  
E--> 2525 ~~87 90~~ 85 90  
2528 <210> SEQ ID NO: 66  
2529 <211> LENGTH: 145  
2530 <212> TYPE: PRT  
2531 <213> ORGANISM: Artificial Sequence  
2533 <220> FEATURE:  
2534 <223> OTHER INFORMATION: hCG beta-subunit with Cys substituted for Ser96 and hFSH  
beta-subunit residues 95-108 for hCG beta-subunit residues 101-108  
2536 <400> SEQUENCE: 66  
2538 Ser Lys Glu Pro Leu Arg Pro Arg Cys Arg Pro Ile Asn Ala Thr Leu  
2539 1 5 10 15  
2542 Ala Val Glu Lys Glu Gly Cys Pro Val Cys Ile Thr Val Asn Thr Thr  
2543 20 25 30  
2546 Ile Cys Ala Gly Tyr Cys Pro Thr Met Thr Arg Val Leu Gln Gly Val  
2547 35 40 45  
2550 Leu Pro Ala Leu Pro Gln Val Val Cys Asn Tyr Arg Asp Val Arg Phe  
2551 50 55 60  
2554 Glu Ser Ile Arg Leu Pro Gly Cys Pro Arg Gly Val Pro Asn Val Val  
2555 65 70 75 80  
2558 Ser Tyr Ala Val Ala Leu Ser Cys Gln Cys Ala Leu Cys Arg Arg Cys  
2559 85 90 95

number the  
amino acids  
under every 5  
amino acids

last sequence in submitted file

P. 10

## RAW SEQUENCE LISTING

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:10

Input Set : A:\Sequen~3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

2562	Thr	Thr	Asp	Cys	Thr	Val	Arg	Gly	Leu	Gly	Pro	Ser	Tyr	Cys	Ser	Phe
2563				100					105					110		
2566	Gly	Glu	Phe	Gln	Asp	Ser	Ser	Ser	Ser	Lys	Ala	Pro	Pro	Pro	Ser	Leu
2567				115					120					125		
2570	Pro	Ser	Pro	Ser	Arg	Leu	Pro	Gly	Pro	Ser	Asp	Thr	Pro	Ile	Leu	Pro
E--> 2571				130					135					140		

*insert*  
*hard*  
*return*  
*145E*

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/797,553B

DATE: 03/10/2005  
TIME: 09:59:11

Input Set : A:\Sequen~3.prj  
Output Set: N:\CRF4\03102005\J797553B.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces. ,

Seq#:24; Line(s) 785  
Seq#:57; Line(s) 2173  
Seq#:65; Line(s) 2505  
Seq#:66; Line(s) 2534

10/797,553B 12

SEQUENCE LISTING

<110> Moyle, William R.  
Xing, Yongna

<120> Protein Knobs

<130> 268/279-RWJ-01-40

<1507> <140> 60/345,283

<1517> <141> 2001-11-08

*These are prior data.*

10/292,553B 13

<210> 8  
<211> 92  
<212> PRT  
<213> Artificial Sequence

Gly is at location 22

<220>  
<223> hCG alpha-subunit with Cys substituted for Leu22

<400> 8

Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Cys Gln Glu Asn Pro  
1 5 10 15

Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
20 25 30

10/797,553B

14

<210> 28  
<211> 92  
<212> PRT  
<213> Artificial Sequence

Ser is at 64

<220>  
<223> hCG alpha-subunit with Cys substituted for Ser64

<400> 28

Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
1 5 10 15

Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
20 25 30

Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
35 40 45

Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Cys Ser  
50 55 60



10/797,553B

15

<210> 44  
<211> 408  
<212> PRT  
<213> Artificial Sequence

<220>

<223> hCGbeta,S138C-betaLA(long), beta-lactamase fused to the carboxy-  
terminal end of hCGb,S138C

insert a  
space after  
"carboxy."  
move "te" down

**VERIFICATION SUMMARY**

DATE: 03/10/2005

PATENT APPLICATION: US/10/797,553B

TIME: 09:59:11

Input Set : A:\Sequen-3.prj

Output Set: N:\CRF4\03102005\J797553B.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:786 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:24  
M:332 Repeated in SeqNo=24  
L:806 M:252 E: No. of Seq. differs, <211> LENGTH:Input:92 Found:184 SEQ:24  
L:1223 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:36  
L:1269 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:37  
L:1317 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:38  
L:1366 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:39  
L:1768 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:45  
L:1906 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:48  
L:2162 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:56  
L:2162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0  
L:2162 M:333 E: Wrong sequence grouping, Amino acids not in groups!  
L:2163 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:56  
L:2170 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:57  
L:2193 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:57  
M:332 Repeated in SeqNo=57  
L:2196 M:252 E: No. of Seq. differs, <211> LENGTH:Input:92 Found:107 SEQ:57  
L:2241 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:58  
L:2287 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:59  
L:2333 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:60  
L:2378 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:61  
L:2423 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:62  
L:2468 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:63  
L:2473 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:64  
L:2496 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:64  
L:2502 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:65  
L:2525 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:65  
L:2571 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:66  
L:13 M:203 E: No. of Seq. differs, <160> Number Of Sequences:Input (56) Counted (66)